NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority under 35 U.S.C. § 120 to PCT International Application No. PCT/FI03/00110, which has an international filing date of February 12, 2003, which designated the United States, which claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Application No. 60/355,782, filed on February 12, 2002. This application is also a Continuation-in-Part of copending Application No. 10/364,505, filed on February 12, 2003, which claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Application No. 60/355,782, filed on February 12, 2002. The entire contents of all of the above application are hereby incorporated by reference.

The present invention relates to a novel human gene functionally related to dyslexia, especially variant forms (e.g. alleles) thereof that predispose an individual to develop dyslexia. Thus, this invention also relates to the polymorphism of said gene as well as diagnostic methods and materials for analysing allelic variation in said gene. This invention also provides polypeptides encoded by said gene and antibodies binding to said polypeptides. The materials of the invention can be used to study the brain processes such as reading, phonological processing, rapid naming and verbal short term memory.

BACKGROUND OF THE INVENTION

Dyslexia, or specific reading disability, is the most common childhood learning disorder. It is estimated that about 3-10% of people have specific difficulties in reading, despite adequate intelligence, education and social environment. Several different theories have been put forth to account for the diverse symptoms seen in dyslexic subjects. At present, it is thought that dyslexia is primarily a phonological deficit, emphasizing the linguistic basis of this condition (1-3). However, it is possible that dyslexia is not specific to language. Rather, it may result from a deficit in processing fast temporal data, be it visual or auditory. This temporal processing deficit would, consequently, manifest itself primarily as dyslexia (4).

Available evidence suggests that dyslexia is a neurological disorder with a genetic basis. Functional brain imaging studies have illustrated that dyslexia has universal neurobiological correlates (5). There is extensive evidence of genetic factors

which contribute to dyslexia. There are significant differences, however, in the heritability of different components of dyslexia (6). Linkage and association studies have pinpointed several loci for dyslexia. In particular, two loci have been promising. *DYX1* in chromosome 15q21 was the first locus to be associated with dyslexia (7), and the results have been replicated in three independent studies thereafter (8-10). The presence of a second dyslexia locus, *DYX2*, in chromosome 6p21 has also been established (11).

We have previously reported a translocation t(2;15)(q11;q21) which segregates with dyslexia (12). In the present invention, we have cloned the breakpoint and narrowed down the breakpoint interval within a 3229 bp region with Southern hybridization. This region contains a 301 bp AT rich sequence. Considering that AT rich repeats are known to occur at many chromosomal rearrangement sites (13), the 301 bp AT rich sequence is likely to be the exact breakpoint site. Furthermore, we have unexpectedly discovered and characterized a novel gene residing in the breakpoint region, which gene we named *DYXC1* and which is causally correlated with dyslexia.

A candidate gene for developmental dyslexia is disclosed in Taipale *et al.* (Proceedings of the National Academic of Sciences 100:11553-11558, 2003), which is incorporated herein by reference.

SUMMARY OF THE INVENTION

The present invention describes a novel human gene, *DYXC1*, which is causally correlated with dyslexia. The coding sequence of *DYXC1* is 1260 bp in length (SEQ ID NO: 1), and it encodes a predicted protein of 420 amino acids (SEQ ID NO: 3). *DYXC1* is expressed in several tissues, most abundantly in brain, lung, kidney and testis. DYXC1 protein resides in cell nuclei, and in brain, it localizes to a subset of cortical neurons and glial cells. DYXC1 protein appears rapidly upregulated and translocated after brain ischemia. The predicted 420 amino-acid protein contains three C-terminal tetratricopeptide repeat (TPR) domains, thought to mediate protein-protein interactions. Besides these domains, it bears no similarity to known proteins. Transfection and immunofluorescence studies indicate that DYXC1 is a nuclear protein.

The coding sequence of *DYXC1* was predicted from the genomic sequence of BAC clones RP11-178D12 and CTD-2137J4. The length of *DYXC1* mRNA is 1993 bp (SEQ ID NO: 2), and it encodes a predicted protein of 420 amino acids. *DYXC1* consists

of 10 exons spanning approximately 78 kb of genomic DNA (Fig. 1D). The start codon (AUG) of *DYXC1* is located 369 bp from the predicted transcription initiation site in exon 2. Putative promoter of *DYXC1* has a TATA box (TATAAAT) at position –31.

In one aspect, the invention features isolated *DYXC1* nucleic acid molecules having the sequence of SEQ ID NO:1 or a complement thereof; homologs and variants thereof as well as fragments thereof. In a preferred embodiment the isolated *DYXC1* nucleic acid is mammalian. In an even more preferred embodiment the isolated *DYXC1* nucleic acid is from a primate, most preferably the *DYXC1* nucleic acid is human. The invention features also vectors comprising the disclosed nucleic acid as well as host cells for the expression or amplification of such vectors.

In addition, we have characterized in this invention five single nucleotide polymorphisms (SNPs) in *DYXC1* mRNA. One sequence variant (1249G→T) introduces a premature stop codon and is inherited with dyslexia in a three-generation family. The frequency of the polymorphism is significantly (p=0.0278) elevated in dyslexic subjects, compared to control samples. The polymorphism truncates the predicted DYXC1 protein by four amino acids, suggesting that it is a functional SNP. Thus, in another aspect, the invention features nucleic acids comprising at least one single nucleotide polymorphism in any one of the following positions as defined by SEQ ID NO:1: position 4 (C preferably to T), 572 (G preferably to A), 1249 (G preferably to T), 1259 (C preferably to G); and SEQ ID NO:2: position 205 (C preferably to T).

The invention further provides polypeptides encoded by *DYXC1* gene or allelic variants thereof and antibodies binding to said polypeptides. The invention also relates to diagnostic methods, kits and materials for analysing allelic variation in *DYXC1* gene and its cellular function.

DESCRIPTION OF THE DRAWINGS

Figures 1A, 1B, 1C and **1D. 1A**, Pedigree of the studied family. Black fill denotes translocation, grey fill dyslexia. **1B**, Fluorescent *in situ* hybridization with BAC clone 178D12 as a probe, showing hybridisation signals in chromosomes 15, der(15), and der(2). **1C**, Southern hybridization with a probe derived from 178D12 shows genomic rearrangement with six restriction enzymes in the studied sample (T) compared to the control sample (C). **1D**, Physical map of the breakpoint region, including *DYXC1*

(black) and an intronic pseudogene (white), drawn to scale. Black triangle illustrates the Southern hybridisation probe position, grey bar denotes the breakpoint interval.

Figures 2A, 2B, 2C, 2D and **2E. 2A**, Comparison of the protein sequences of human *DYXC1* and mouse *mdyxc1*. The SNPs found in this invention are marked with a circumflex accent, and the three TPR domains are marked with asteriks. **2B**, RT-PCR from human multiple tissue cDNA panels I and II (Clontech). Lanes: 1)λ/ϕX174 size marker, 2) heart, 3) brain, 4) placenta, 5) lung, 6) liver, 7) skeletal muscle, 8) kidney, 9) pancreas, 10) spleen, 11) thymus, 12) prostate, 13) testis, 14) ovary, 15) small intestine, 16) colon, and 17) leukocyte. **2C** and **2D**, *DYXC1* Northern blot from Multiple Tissue Northern (MTN) Blot panels I and II (Clontech). Lanes in fig. **2C**: 1) heart, 2) brain, 3) placenta, 4) lung, 5) liver, 6) skeletal muscle, 7) kidney, 8) pancreas; Lanes in fig. **2D**: 9) spleen, 10) thymus, 11) prostate, 12) testis, 13) ovary, 14) small intestine, 15) colon, and 16) leukocyte. **2E**, Cellular localization of *DYXC1* protein. Cos-1 cells transfected with *DYXC1*-V5 fusion construct were stained with monoclonal mouse α-V5 antibody and FITC-conjugated α-mouse-IgG (grey). DAPI stained nuclei are shown in light grey.

Figure 3. Pedigree of the family in which two *DYXC1* polymorphisms are transmitted with dyslexia. Alleles for four loci are shown below individuals. Black rectangle indicates the haplotype transmitted with dyslexia in this family.

Figure 4. Immunostaining patterns for DYXC1 observed in normal human brain tissue from an individual died immediately after sudden cardiac arrest. A, Characteristic immunoreactivity in cortical brain tissue demonstrating a variable density of nuclear expression in a minority of neurons. (Original magnification x30). B, Typical staining result in white matter, where also a fraction of cell nuclei are densely positive in contrast to clearly negative adjacent cell nuclei. (x40). C, Positive neuronal nucleus adjacent to negative glial cell nuclei (x100). D, Negative neuronal cell body adjacent to neighboring small cells representing probably glial cells (x100). E, High magnification illustration of a large pyramidal neuron expressing clearly intranuclear localization of DYXC1 protein (x100). F, Typical view of an adjacent tissue section stained with preimmune (control) serum indicating the lack of non-specific staining in neuronal and glial cell nuclei (x30).

Figure 5. Immunostaining patterns for DYXC1 observed in human ischemic brain tissue of three victims of acute ischemic stroke. A, Typical immunostaining result in a subject deceased at 23 h after the onset of stroke symptoms. Note the characteristically cytoplasmic staining pattern present in a subset of ischemic neurons. Compare to Fig. 4 from non-ischemic brain tissue, where typically nuclear expression was found. (x40). B, Same neuronal population identified from adjacent section stained with preimmune (control) serum (x40). C, Typical immunostaining pattern in ischemic brain tissue area of another subject deceased 26 h after the onset of stroke symptoms. Note the increased fraction of pyknotic neurons expressing dense, predominantly nuclear, DYXC1 immunoreactivity. (x20). D, Same neuronal population identified from adjacent section stained with preimmune (control) serum (x20). E, Immunoreactivity for DYXC1 observed in a more advanced, vacuolized, stage of tissue ischemia demonstrating faint expression also in neuronal processes. This subject died 60 h after the onset of stroke symptoms. (x40). F, Same tissue area identified from adjacent section stained with preimmune (control) serum (x40).

DETAILED DESCRIPTION OF THE INVENTION

This invention is based on the discovery and characterization of a novel human gene termed *DYXC1*. The human *DYXC1* gene is 1260 bp in length (SEQ ID NO: 1) and it encodes a 420-amino acid residue protein (SEQ ID NO: 3). The cDNA of total *DYXC1* mRNA (SEQ ID NO:2) has been deposited in GenBank with accession number AF337549. *DYXC1* maps to human chromosome 15q21. The present invention shows that previously reported balanced translocation breakpoint t(2;15)(q11;q21) segregating coincidentally with developmental dyslexia is located in *DYXC1* thus indicating that *DYXC1* is linked to dyslexia. In addition, it was unexpectedly discovered in the present invention that point mutations, i.e. SNPs, in *DYXC1* segregate with the susceptibility to develop dyslexia.

The present invention provides *DYXC1* nucleic acids, homologs thereof and fragments thereof. The human *DYXC1* cDNA sequence is disclosed in SEQ ID NO: 1. Preferred homologs, such as chimpanzee (SEQ ID NO:13), pygmy chimpanzee (SEQ ID NO:19),

gorilla (SEQ ID NO:15), orangutan (SEQ ID NO:17) and mouse *dyxc1* (SEQ ID NO: 4), have a sequence at least about 79% homologous with a nucleotide sequence of SEQ ID NO: 1. In a preferred embodiment, the *DYXC1* nucleic acid is from a mammal, e.g. a mouse, primate or human. In another preferred embodiment the nucleic acid has the sequence of SEQ ID NO: 1 a complement thereof or a fragment thereof. In one embodiment of the invention the fragment disclosed can be a primer or probe, which is capable to hybridise specifically to the *DYXC1* nucleic acids described herein. The preparation and modification of primers and probes capable of binding to a known nucleic acid are well-established techniques in the art (see, for example, Current Protocols in Molecular Biology, eds. Ausubel et al, John Wiley & Sons:1992). Generally, a primer or a probe is a substantially purified oligonucleotide being 12 to 60 nucleotides long, preferably 16 to 40 nucleotides. A primer or probe need not reflect the exact sequence of a template, i.e. a target nucleic acid, but must be sufficiently complementary to hybridise with the template under stringent conditions.

The invention also involves nucleotide sequence variants capable of encoding *DYXC1* polypeptides. Such variants include sequences that differ from the disclosed *DYXC1* nucleic acids by one or more nucleotide substitutions, additions or deletions, such as allelic variants. Said nucleotide substitutions may also arise due to the degeneracy of the genetic code. The nucleic acids of the invention can also be described as capable of hybridising under stringent conditions to the nucleic acid sequence of SEQ ID NO: 1 or 2 or a complement thereof. Such stringent DNA hybridisation conditions are well-known in the art, e.g. 6×NaCl/sodium citrate (SSC) at about 45 °C is applied for a hybridisation step, followed by a wash of 2×SSC at 50 °C or , e.g., alternatively hybridization at 42 °C in 5×SSC, 20 mM NaPO4, pH 6.8, 50% formamide; and washing at 42 °C in 0.2×SSC. Those skilled in the art understand that it is desirable to vary these conditions empirically based on the length and the GC nucleotide base content of the sequences to be hybridised, and that formulas for determining such variation exist (See, for example, Sambrook et al,"Molecular Cloning: A Laboratory Manual", Second Edition, pages 9.47-9.51, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1989)). Nucleic acids of the invention, fragments thereof and variants thereof with sufficient similarity to the non-coding strand of said nucleic acids to hybridise thereto under stringent conditions are useful for identifying, purifying, and

isolating nucleic acids encoding other, non-human, mammalian forms of *DYXC1*. Thus, such polynucleotide fragments and variants are intended as aspects of the invention.

The present invention also provides plasmids and vectors encoding an *DYXC1* polypeptide, which constructs can be used in the expression of said *DYXC1* polypeptide in or from a host cell. The selecting of a suitable plasmid or vector for a certain use is within the abilities of a skilled artisan. As the host cell may be any prokaryotic or eukaryotic cell, a plasmid or vector encoding an *DYXC1* polypeptide can be used to the production of said *DYXC1* polypeptide as a recombinant protein via microbial or eukaryotic cellular processes. Typically, said plasmids and vectors comprises a ligated nucleic acid encoding a recombinant protein, said nucleic acid operably linked to at least one transcriptional regulatory sequence (See, for example, Sambrook et al,"Molecular Cloning: A Laboratory Manual", Second Edition, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1989)).

The present invention further describes the characterization of single nucleotide polymorphisms (SNPs) in human *DYXC1* gene. SNPs can be used in mapping the human genome and, when a SNP is linked with a disease or condition, to clarify genetic basis of the disease or condition, in this particular case, at least of dyslexia. In this invention we have characterized five single nucleotide polymorphisms (SNPs) in *DYXC1* mRNA (SEQ ID NOS:1 and 2). Accordingly, the present invention provides an *DYXC1* nucleic acid comprising a SNP in any one of the following positions in the nucleic acid sequence of SEQ ID NO: 1: position 4; 572; 1249 or 1259. Allelic variation at position 4 consists of a single base substitution from C preferably to T. Allelic variation at position 572 consists of a single base substitution from G preferably to A. Allelic variation at position 1249 consists of a single base substitution from G preferably to T. Allelic variation at position 1259 consists of a single base substitution from C preferably to G. The present invention also describes a SNP in *DYXC1* mRNA at position –164 outside the coding sequence of *DYXC1*. This position corresponds to position 205 set forth in SEQ ID NO:2. Allelic variation at position -164 consists of a single base substitution from C preferably to T.

The SNP variant of *DYXC1*, wherein the single base substitution is at position 1249 (G→T), introduces a premature stop codon and is inherited with dyslexia in a three-generation family. The frequency of the polymorphism is significantly (p=0.0278) elevated in dyslexic subjects, compared to control samples as shown in Examples. The polymorphism truncates the predicted DYXC1 protein by four amino acids, suggesting that it is a functional SNP.

Further, new polymorphic gene regions in *DYXC1* nucleic acids can be identified by determining the *DYXC1* nucleic acid sequences in population of inviduals. If new polymorphic region (e.g. SNP) is found, then the link with a specific disease can be determined by studying specific populations of individuals, such as dyslexics. A polymorphic site or region may be located in any part of a gene, e.g., exons, introns and promoter region.

The present invention makes available DYXC1 polypeptides. Such polypeptides can be recombinant proteins produced by, e.g., the host cells described hereinabove, said recombinant proteins being isolated from other cellular proteins. Preferably, said polypeptides have an amino acid which is at least about 78% identical or homologous to human DYXC1 protein of sequence set forth in SEQ ID NO: 3. In a preferred embodiment, an DYXC1 polypeptide of the present invention is mammalian, e.g. murine or human, DYXC1 protein. In addition, the present invention provides splice variants of DYXC1 protein.

An DYXC1 polypeptide of the invention can also be used as an antigen to produce antibodies. Techniques of preparing antisera, poly- or monoclonal antibodies are well-known protocols in the art (see, for example, Antibodies: A laboratory Manual, eds. Harlow and Lane, Cold Spring Harbor Laboratory Press: 1988). Thus, the present invention makes available DYXC1 specific antibodies. Especially, the antibodies of the invention can be labeled with a detectable label and used in the determination of the presence of DYXC1 polypeptides in a sample, e.g. for diagnosis of dyslexia.

The present invention further provides means for prognostic or diagnostic assays for determining if a subject has or is likely to develop dyslexia, which is associated with the

variation or dysfunction of *DYXC1*. Basically, such assays comprise a detection step, wherein the presence or absence of a genetic alteration or defect in *DYXC1* is determined in a biological sample from the subject. Said detection step can be performed, e.g., by methods involving sequence analysis, nucleic acid hybridisation, primer extension, restriction enzyme site mapping or antibody binding. These methods are well-known in the art (see, for example, Current Protocols in Molecular Biology, eds. Ausubel et al, John Wiley & Sons:1992).

In particular, the present invention is directed to a method of determining the presence or absence of an *DYXC1* SNP of the invention in a biological sample from a human for diagnostics of dyslexia or for assessing the predisposition of an individual to dyslexia. Said method comprises determining the sequence of the nucleic acid of a human at one or more positions 4, 572, 1249 and 1259 in the *DYXC1* gene or mRNA as defined in SEQ ID NO:1 and position 205 as defined by SEQ ID NO:2 and determining the status of the human by reference to polymorphism in *DYXC1* gene. In a preferred embodiment the sample is contacted with oligonucleotide primers so that the nucleic acid region containing the potential single nucleotide polymorphism is amplified by polymerase chain reaction prior to determining the sequence. The final results can be obtained by using a method selected from, e.g., allele specific nucleic acid amplification, allele specific nucleic acid hybridisation, oligonucleotide ligation assay or restriction fragment length polymorphism (RFLP). These methods are well-known for a skilled person of the art (see, for example, Current Protocols in Molecular Biology, eds. Ausubel et al, John Wiley & Sons:1992, or Landegren et al, "Reading Bits of Genetic Information: Methods for Single-Nucleotide Polymorphism Analysis", Genome Research 8:769-776).

The invention also features diagnostic or prognostic kits for use in detecting the presence of *DYXC1* SNP in a biological sample. The kit provides means for the diagnostics of dyslexia or for assessing the predisposition of an individual to dyslexia mediated by variation or dysfunction of *DYXC1*. The kit can comprise a labeled compound capable of detecting *DYXC1* polypeptide or nucleic acid (e.g. mRNA) in a biological sample. The kit can also comprise nucleic acid primers or probes capable of hybridising specifically to at least of portion of an *DYXC1* gene or allelic variant

thereof. The kit can be packaged in a suitable container and preferably it contains instructions for using the kit.

It is also realised in the present invention that transgenic non-human animals, such as transgenic mice, which include a heterologous form of an *DYXC1* gene, can be designed and produced utilising the disclosure presented herein (see, for example, Manipulating the Mouse Embryo: A laboratory Manual, eds. Hogan et al, Cold Spring Harbor Laboratory Press, 1986). Such transgenic animals can be useful as animal models for studying, e.g., the function of *DYXC1* gene and alleles thereof, or for expressing recombinant DYXC1 polypeptides.

A further embodiment of the present invention is a method for identifying a mutant *DYXC1* nucleotide sequence in a suspected mutant *DYXC1* allele which comprises comparing the nucleotide sequence of the suspected mutant *DYXC1* allele with a wild-type *DYXC1* nucleotide sequence or a part thereof, wherein a difference between the suspected mutant and the wild-type sequence identifies a mutant *DYXC1* nucleotide sequence. In said method the sequence of said suspected mutant *DYXC1* allele can be compared with the sequence of one or more wild-type *DYXC1* gene sequences selected from the sequences set forth in SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10 and wild-type allelic variants thereof. For the screening of new point mutations, deletion mutations and insertion mutations in *DYXC1* a plentiful of techniques well-known for a skilled artisan can be utilised, such as methods involving sequence analysis, nucleic acid hybridisation, primer extension, restriction enzyme site mapping and particularly methods described below in Experimental Section and Materials and Methods.

Screening assays

The subject methods include screens for agents which modulate the activity of DYXC1 gene or protein. The methods are amenable to automated, cost-effective high throughput screening of chemical libraries for lead compounds. Identified reagents find use in the pharmaceutical industries for animal and human trials; for example, the reagents may be derivatized and rescreened in vitro and in vivo assays to optimize activity and minimize

toxicity for pharmaceutical development. More specifically, identified reagents may find use in the treatment of dyslexia or brain ischemia.

The invention further provides methods (also referred to herein as "screening assays") for identifying modulators, i.e., candidate or test compounds or agents (e.g., peptides, peptidomimetics, peptoids, small molecules or other drugs) which bind to DYXC1 protein, have a stimulatory or inhibitory effect on, for example, DYXC1 expression or DYXC1 activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of a DYXC1 substrate. Compounds thus identified can be used to modulate the activity of DYXC1 in a therapeutic protocol, to elaborate the biological function of the DYXC1, or to identify compounds that disrupt normal DYXC1 activity. The preferred DYXC1 used in this embodiment are human, primate or mouse DYXC1 of the present invention.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a DYXC1 protein or polypeptide or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a DYXC1 protein or polypeptide or biologically active portion thereof.

The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries [libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive] (see, e.g., Zuckermann, R. N. et al. J. Med. Chem. 1994, 37: 2678-85); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the `one-bead one-compound` library method; and synthetic library methods using affinity chromatography selection. The biological library and peptoid library approaches are limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K. S. (1997) Anticancer Drug Des. 12:145).

In one embodiment, an assay is a cell-based assay in which a cell which undergoes a simulated ischaemia is contacted with a test compound and the ability of the test compound to modulate DYXC1 activity is determined. Determining the ability of the test compound to modulate DYXC1 activity can be accomplished by monitoring, for example, cell death, cell growth, cell attachment, and cell chemotaxis. The cell, for example, can be of mammalian origin, e.g., a neuronal cell or a non-neuronal cell. In preferred embodiment, the ability of the test compound to modulate DYXC1 activity is accomplished by monitoring DYXC1 activation with Western blot, immunohistochemical staining using anti DYXC1 antibodies, or fluorometric assays.

Determining the ability of DYXC1 protein or a biologically active fragment thereof, to bind to or interact with an agent can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of DYXC1 protein to bind to or interact with an agent can be accomplished by determining the activity of DYXC1 protein. For example, the activity of DYXC1 can be determined by detecting the induction of a reporter gene (recombinant DYXC1 gene products labelled with detectable marker), or detecting a target-regulated cellular response (i.e., cell attachment, cell adhesion, cell growth, cell death, neurite outgrowth or cell migration).

In yet another embodiment, an assay of the present invention is a cell-free assay in which DYXC1 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to DYXC1 protein or biologically active portion thereof is determined.

Assays for the Detection of the Ability of a Test Compound to Modulate Expression of DYXC1

In another embodiment, modulators of DYXC1 expression are identified in a method wherein a cell is contacted with a candidate compound/agent and the expression of DYXC1 mRNA or protein in the cell is determined. The level of expression of DYXC1 mRNA or protein in the presence of the candidate compound is compared to the level of expression of DYXC1 mRNA or protein in the absence of the candidate compound. The

candidate compound can then be identified as a modulator of DYXC1 expression based on this comparison. For example, when the expression of DYXC1 mRNA or protein is higher (i.e. statistically significantly higher) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of DYXC1 mRNA or protein expression. Alternatively, when expression of DYXC1 mRNA or protein is lower (i.e. statistically significantly lower) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of DYXC1 mRNA or protein expression. The level of DYXC1 mRNA or protein expression in the cells can be determined by methods described herein for detecting DYXC1 mRNA or protein or by methods which a skilled artisan can readily adapt for use in the present invention.

Combination Assays

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of DYXC1 protein can be confirmed in vivo, e.g., in an animal such as an animal model for brain ischemia.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a DYXC1 modulating agent, an antisense DYXC1 nucleic acid molecule, a DYXC1-specific antibody, or a DYXC1-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

The choice of assay format will be based primarily on the nature and type of sensitivity/resistance protein being assayed. A skilled artisan can readily adapt protein activity assays for use in the present invention with the genes identified herein.

One preferred embodiment of the invention is a screening method, wherein a compound that modulates the expression of DYXC1 is identified, the method comprising:

(a) incubating a cell that can express DYXC1 gene with a compound under conditions and for a time sufficient required for the cell to express DYXC1 gene, when the compound is not present;

(b) incubating a control cell under the same conditions and for the same time without the compound;

(c) measuring expression of DYXC1 gene in the cell in the presence of the compound;

(d) measuring expression of DYXC1 gene in the control cell; and

(e) comparing the amount of expression of DYXC1 gene in the presence and absence of the compound, wherein a difference in the level of expression indicates that the compound modulates the expression of DYXC1 gene

Another preferred embodiment of the invention is a method of identifying a compound that modulates DYXC1 activity, the method comprising:

(a) incubating a cell that has said activity with a compound under conditions and for a time sufficient required for the cell to express said activity, when the compound is not present;

(b) incubating a control cell under the same conditions and for the same time without the compound;

(c) measuring said activity in the cell in the presence of the compound;

(d) measuring said activity in the control cell; and

(e) comparing the amount of said acitivity in the presence and absence of the compound, wherein a difference in the level of activity indicates that the compound modulates the activity of said gene

Another preferred emdiment of the invention is a method for affinity purification of a substance that binds to the DYXC1, comprising the following steps: a) contacting a source suspected to contain said substance with an immobilized DYXC1 under conditions whereby said substance to be purified is selectively adsorbed onto the immobilized DYXC1; (b) washing the immobilized DYXC1 and its support to remove non-adsorbed material; and (c) eluting said substance from the immobilized DYXC1 to which they are adsorbed with an elution buffer.

EXPERIMENTAL SECTION

Fine mapping of the translocation breakpoint

The translocation and the phenotypes of the members of the family studied here (Fig. 1A) have been described previously (12). Fluorescent in situ hybridization (FISH) restricted the location of the translocation breakpoint within the BAC clone RP-11-178D12 (AC013355; Fig. 1B). The clone contained two known genes, cell-cycle restoration protein 8 (CPR8) and complementation class B phosphoinositol glycan (PIG-B), in addition to the genes described herein. To further localize the breakpoint, we used amplified non-repetitive genomic DNA fragments from the BAC clone RPCI-11-178D12 as probes in Southern hybridization. A probe corresponding to nucleotides 102317-102837 of the complete sequence of 178D12 revealed a genomic rearrangement with 6 different restriction enzymes (Fig. 1C). Thus, we could pinpoint the breakpoint to a region of 3229 bp, limited by the restriction sites for PstI and HindIII (Fig. 1D). The breakpoint region includes exons 8 and 9 of a novel gene, *DYXC1* (see below).

Characterization of *DYXC1*

The coding sequence of *DYXC1* was predicted *in silico* from the genomic sequence of BAC clones RP11-178D12 and CTD-2137J4. Exon-intron boundaries were confirmed with RT-PCR. The length of *DYXC1* mRNA, obtained by RT-PCR, is 1993 bp, and it encodes a predicted protein of 420 amino acids. *DYXC1* consists of 10 exons spanning approximately 78 kb of genomic DNA (Fig. 1D). Three promoter prediction programs (see below) identified a promoter precisely before the 5' end of the RT-PCR obtained

mRNA, suggesting that the cloned mRNA is nearly complete. The start codon (AUG) of *DYXC1* is located 369 bp from the predicted transcription initiation site in exon 2. Putative promoter of *DYXC1* has a TATA box (TATAAAT) at position –31.

Database searches revealed several mouse ESTs homologous to the human *DYXC1* mRNA. We could thus construct the mouse *mDYXC1 in silico* by connecting the overlapping EST clones. The *mDYXC1* mRNA (SEQ ID NO: 4) encodes a 421-residue protein (SEQ ID NO: 5) that is 78% identical with the human DYXC1 (Fig. 2A). The human DYXC1 protein does not have any significant homologies to other known proteins. It has, however, three C-terminal tetratricopeptide repeat (TPR) domains, corresponding to amino acids 290-323, 324-357, and 366-399. These TPR domains are thought to mediate protein-protein interactions (14, 15).

The human *DYXC1* mRNA appears to exist in several different splice forms: exons 9 and 2 can be omitted, and there is an alternative acceptor splice site in intron 2. All these arrangements, however, alter the reading frame, leading to truncated protein products. *DYXC1* mRNA can be found in several tissues. It is most abundantly expressed in brain, lung, kidney and testis (Fig. 2B). Northern blot of human adult kidney tissue revealed an approximately 2 kb transcript, corresponding to the predicted size of *DYXC1* mRNA (Fig. 2C).

The cellular localization of DYXC1 in transfected monkey kidney COS-1 cells was studied using immunofluorescence. The full-length *DYXC1* cDNA was cloned into a mammalian expression vector containing a C-terminal V5 epitope and a polyhistidine tail. DYXC1-V5/His fusion protein showed a staining pattern similar to DAPI staining, suggesting that DYXC1 is a nuclear protein (Fig. 2E).

Association analysis of *DYXC1* SNPs

The present invention also describes single nucleotide polymorphism in *DYXC1*. In order to find polymorphisms in *DYXC1*, we screened the *DYXC1* cDNA from 57 dyslexic individuals from 22 unrelated families with single-stranded conformation polymorphism (SSCP) analysis to characterize sequence variants. As a control, we screened 91 anonymous blood donors from Turku and Kuopio, and 15 non-dyslexic

subjects from the 22 dyslexia families. In this way, we found three SNPs. Two of the SNPs (4C→T, 572G→A) were in the coding region, whereas a third one (-164C→T) resided in the 5' untranslated region (Table 1). Both the SNPs in the coding region resulted in amino acid substitutions.

Table 1. Frequency of single nucleotide polymorphisms in dyslexic subjects and controls.

SNP	Codon change	Amino acid position	Frequency				p-value
			Dyslexia	n	Control	n	
-164C→T*	-	-	0.0517	58	0.0286	105	0.2219
-3G→A**	-	-	0.083	54	0.031	113	0.0020
-2G→A***	-	-	0.009	54	0	113	0.7150
4C→T	Pro → Ser	2	0.0172	58	0	105	0.1259
271G→A	Val → Ile	90	0.019	54	0.053	113	0.9248
572G→A	Gly → Glu	191	0.4811	53	0.5202	99	0.7792
1249G→T	Glu → STOP	417	0.1228	57	0.0545	101	0.0278
1259C→G	Ser → Cys	420	0,0789	57	0,0769	104	0,9482

* corresponds to position 205 set forth in SEQ ID NO:2

** corresponds to position 366 set forth in SEQ ID NO:2

*** corresponds to position 367 set forth in SEQ ID NO:2

4C→T, a nonconservative substitution of proline-2 to serine-2, was found from two dyslexic individuals (a father and a son), but not in any of the control subjects. Further examination showed that this alteration did not segregate with dyslexia in the extended pedigree. Likewise, the frequency of 572G→A did not differ significantly between the two groups. The third SNP, -164C→T was found in 6 dyslexic individuals from three families and in 5 control subjects. In one three-generation family, the T allele segregated with dyslexia (Fig. 3). In the other two families, there was no consistent pattern of inheritance with dyslexia.

To search for additional SNPs, we sequenced the whole coding region of *DYXC1* from an individual carrying the T allele in the family presented in Figure 3. We found a G to

T transversion at position 1249 of the *DYXC1* mRNA, which results in a substitution of a glutamic acid for an ochre stop codon at amino acid position 417. The appearance of a stop codon leads to the deletion of the C-terminal tetrapeptide Glu-Leu-Lys-Ser. In the family of Figure 3, 1249G→T was transmitted in the same chromosome as -164C→T, thus segregating with dyslexia.

Screening of all 57 dyslexic subjects for 1249G→T showed that the SNP is relatively common with a frequency of 0.123 (14 /114 chromosomes). In the control group, the frequency was only 0.055 (10/174 chromosomes in blood donor samples, 1/28 chromosomes in control subjects of dyslexia families). All the control subjects were heterozygous for the SNP, whereas there was one dyslexic subject homozygous for the SNP. In conclusion, the frequency of 1249G→T is significantly (p=0.0278) higher in dyslexic individuals.

In further studies we found three more SNPs. One of the SNPs (271G→A) was in the coding region and resulted in amino acid substitution, whereas two (-3G→A and -2G→A) resided in the 5' untranslated region (Table 1). SNP -3G→A showed significant association with dyslexia ($P = 0.006$). SNP -3G→A is located in the binding sequence of the transcription factors Elk-1, HSTF, and TFII-I. Elk-1 is a transcriptional activator expressed in rat brain neurons and its activation has been associated with learning in rats. The -3G→A SNP affects also the Kozak sequence near the translation initiation site.

Primate genes

The nonhuman primates chimpanzee, pygmy chimpanzee, gorilla and orangutan were sequenced for the genomic sequences corresponding to human exons and differed for 3, 2, 5 and 6 amino acids (0.7%, 0.5%, 1.2% and 1.4% of residues), respectively (Table 2).

Table 2. Comparison of *DYXC1* cDNA between human and four nonhuman primates. Nucleic acid and amino acid changes are shown for each exon of *DYXC1* in comparison to the human sequence; + indicates the presence of a change in a nonhuman species.

Exon	Nucleic acid change	Amino acid change	Chimpanzee	Pygmy chimpanzee	Gorilla	Orangutan
1	none					
2	6T>C		+	+	+	+
2	47C>T	A16>V			+	
2	48G>C		+			
2	107C>T	T36>M	+			
3	none					
4	284T>C	M95>T				+
4	384C>T		+	+	+	+
5	473C>A	A158>E	+	+	+	+
5	516A>C	Q172>H				+
5	520A>G	K174>E			+	
5	540A>G				+	
5	572G>A	G191>E	+	+	+	+
5	583A>T	I195>L			+	+
5	591T>C					+
5	611T>C	L204>S				+
5	624T>C					+
6	639G>A					+
7	789A>G			+		
8	909G>A		+	+	+	+
9	none					
10	none					

Remarkably, the DYXC1 protein differs from its pygmy chimpanzee and chimpanzee counterparts at 2-3 amino acids, but from gorilla and orangutan at 5-6 residues. For comparison, the level of coding divergence is higher for DYXC1 than that observed for FOXP2, the product of a gene implicated in a speech and language disorder (26). Thus, as suggested for *FOXP2*, the *DYXC1* gene may reveal important evolutionary differences related to brain functions between the primates.

Expression of DYXC1 protein in human brain

Light microscopy of normal human brain sections revealed a strikingly nuclear expression pattern for DYXC1 immunoreactivity, consistent with the transfection results. Both in cortical neuronal cell populations and white matter glial cells, a minority of cell nuclei expressed DYXC1 immunoreactivity (Fig. 4A,B,E) . Characteristically, even neighbouring, identical appearing cells had different expression (Fig. 4C,D), which, together with the lack of staining obtained with preimmune serum (Fig. 4F), supports the specificity of the observed immunoreactivity.

We studied DYXC1 immunoreactivity also in individuals died soon after the onset of acute ischemic stroke (Fig. 5). In contrast to the typically nuclear expression in the normal brain, also cytoplasmic expression was observed in ischemic brain areas (Fig. 5A). In cortical areas representing early ischemic morphology, the fraction of positive cell nuclei or cytoplasms appeared increased (Fig. 5C) as compared to non-ischemic brain or contralateral hemispheres. In most ischemic sections studied, also structures corresponding to neuronal processes were frequently found to be immunoreactive for DYXC1 (Fig. 5E). Since only a limited number of subjects with short post-ischemic intervals were available, quantitative or statistical analysis of expression was not attempted.

Observations from both normal and ischemic human brains demonstrated that DYXC1 protein is expressed in a fraction of human glial and neuronal cells. Since only a fraction of neighbouring, identical appearing cells expressed DYXC1 immunoreactivity, we suggest that DYXC1 does not represent a structural house-keeping element of brain

cells. Instead, it may relate to the functional state of the cells. Examination of DYXC1 expression in ischemic brain tissue suggested that it is involved dynamically in the functional cell state changing in the face of metabolic challenge. Since the topographic expression in the ischemic tissue included cytoplasms and even neuronal processes, DYXC1 may also have extranuclear functions, and may be rapidly translocated from nuclei to cytoplasms when the cell undergoes metabolic stress. These results from human cerebral ischemia warrant more systematic studies on the role of DYXC1 in cell stress and ischemia as well as dyslexia.

MATERIALS AND METHODS

Ascertainment of patients and psychological assessment

Patients were selected among families of dyslexic children from the Department of Pediatric Neurology at the Hospital for Children and Adolescents, University of Helsinki, Finland and from the Association of Learning Disabled Individuals of Helsinki (HERO), participating in the genetic study of dyslexia. The study was approved by the Ethical Committee of The Children's Castle Hospital and informed consent was obtained from the participants. Pedigrees of two families are shown in Fig. 1A and Fig. 3. The diagnosis and degree of dyslexia was determined by Finnish reading and spelling tests designed for children (16) and adults (17). The intelligence quotient (IQ) was determined by WAIS-R (18) or WISC-R (19). Reading-related neurocognitive skills (phonological awareness, rapid naming and verbal short-term memory) were assessed by neuropsychological tests (20-23).

FISH and Southern blotting.

RPCI-11 BAC clone 178D12 (Genbank accession number AC013355) was used as a probe in fluorescent *in situ* hybridization. The protocol for FISH has been previously described (12). 15 μg of total genomic DNA from an individual carrying the translocation and from an unrelated control person was digested with BamHI, EcoRI, HindIII, BsaAI, PstI, or SphI, and run in 0.7% agarose gel. DNA was transferred to Hybond N+ membrane (Amersham Pharmacia Biotech) with standard alkaline blotting

method. PCR fragments derived from human genomic DNA were TA-cloned into pCR2.1 TOPO-TA vector (Invitrogen, Carlsbad, CA), and insert was removed with EcoRI digestion and gel-purified (Qiagen, Venlo, The Netherlands). α32P-labeled insert was used as a probe in Southern hybridization. Hybridization was performed overnight at 65°C in Church buffer (0.5 M $NaHPO_4$, 1 mM EDTA, 7% SDS, 1% BSA), and the filter was washed in 2 x SSC, 0.05% SDS at 65°C for 1 hour. Filters were autoradiographed with a phosphoimager plate.

Cloning of *DYXC1* and sequence analysis

Novel genes in the sequence of clone 178D12 were predicted *in silico* with Genscan (24) and Fgenes software. Predicted genes were confirmed by sequencing RT-PCR products. *DYXC1* cDNA has been deposited in GenBank with accession number AF337549. Mouse *mDYXC1* was constructed from two overlapping EST sequences (accession numbers BG242087 and AK005832) and verified by comparing it to all available mouse *mDYXC1* EST sequences. cDNA sequences of *mDYXC1* and *hDYXC1* were aligned with ClustalX. The alignment was improved manually, and shaded with BOXSHADE. The secondary structure of the TA rich region was predicted with MFOLD (available at http://bioinfo.math.rpi.edu/~mfold/dna/form1.cgi) with default parameters. The expression of *DYXC1* was analyzed by RT-PCR from Clontech's multiple tissue cDNA panels 1 and 2. RT-PCR was performed in 25 µl volume in the following conditions: 94°C 2' (94°C 1', 68°C 2') x 30, 1 x DyNAzyme buffer with $MgCl_2$ (Finnzymes, Espoo, Finland), 0.2 u DyNAzyme II polymerase (Finnzymes), 15 pmol forward primer GTTGACAGAATGCTGTTCCACGTCG (SEQ ID NO:11), 15 pmol reverse primer CAAGCTGAGGCACGAAGAGCAATGA (SEQ ID NO:12). Promoter region of *DYXC1* was predicted with TSSG and TSSW software at Baylor College of Medicine, available at http://searchlauncher.bcm.tmc.edu/seq-search/gene-search.html, and neural network promoter prediction (NNPP) software at University of California, Berkeley, available at http://www.fruitfly.org/seq_tools/promoter.html. The genomic sequences of nonhuman primates corresponding to all exons were determined by direct sequencing after PCR amplification with human-specific intronic primers (primer sequences are listed in Table 3).

Table 3. Human-specific intronic primers for DYXC1.

Primer Name	Primer Sequence	Primer Length	Product Length	Exon
EKN1-1F	AACAGACTGCCTGGTGCTCT	20	268 bp	exon 1
EKN1-1R	CACACCAAAGTTTGAGAACCACT	23		
EKN1-2.1R	AAGATGAGCCTGTTGCTCGT	20	476 bp	exon 2
EKN1-2.1F	CAAGCAGAGGGTATGGGTCTAC	22		
EKN1-2R	AGAAGCTTCGGACCACACC	19	431 bp	exon 2
EKN1-3F	CGCGTGCTTAATTTGTGTAA	20	299 bp	exon 3
EKN1-3R	TCCCCTACACAATATAGGTGCTT	23		
EKN1-4F	AAAGAAATCTCATCCTGGGTCA	22	327 bp	exon 4
EKN1-4R	GAAAATGCTGAGGAAGTCCAG	21		
EKN1-5F	CAATGGCAAGAGTTTAGAGGTATG	24	456 bp	exon 5
EKN1-5R	TCAATGTGCCAAAACAGTAACC	22		
EKN1-6F	TGTTTAGGATTTGGGGGTGA	20	395 bp	exon 6
EKN1-6R	GGAAATTCTAAAACATATTCATGACG	26		
EKN1-7F	CCACTGGAGGAAGATGGAAA	20	244 bp	exon 7
EKN1-7R	TGTCTTCATACATGATAAAGCTCAT	25		
EKN1-8F	GGTAAGCCATCCTCTTTGTCA	21	337 bp	exon 8
EKN1-8R	TCAACTGAACAGAAAAAGATCATCA	25		
EKN1-9F	CTCCCCAAGTGTTGGGATTA	20	305 bp	exon 9
EKN1-9R	TGGAGTCCTTAAAAGTCACGA	21		
EKN1-10F	GGTACTTGTTCTGAACCATGCTACTA	26	502 bp	exon 10
126403-F	CAAGGGCAAGCTTAATTCAGTAACACA	27		

Single-strand conformational polymorphism analysis (SSCP)

DYXC1 exons were amplified with PCR (primer sequences available from the inventors on request) and digested with suitable enzymes to obtain 100-300 bp fragments. Denaturing gel was run for 16 hours at room temperature with 5 W constant power. Gels were stained with silver according to standard protocols.

SNP analysis

Polymorphisms -164C→T, 4C→T, and 572G→A introduced novel Tsp45I, MnlI, and MboII restriction sites, respectively. Exon-specific PCR products were digested with the appropriate enzyme and run on a 1.5% agarose gel (-164C→T, 4C→T) or on a polyacrylamide gel (572G→A), followed by silver staining. As 1249G→T had no effect on restriction sites, exon 10 was directly sequenced from all subjects. A standard one-tailed Fisher exact test was used to evaluate the statistical significance.

Cellular localization of DYXC1

Full-length *DYXC1* cDNA was cloned into pcDNA3.1/V5-6xHis expression vector (Invitrogen). Monkey kidney COS-1 cell line was transfected with 3 μg of the construct, with FuGENE6 (Roche) as a transfection reagent, according to manufacturer's protocols. Cells were stained with mouse anti-V5 antibody (Invitrogen) and FITC-conjugated goat anti-mouse IgG (Sigma-Aldrich). Nuclei were stained with DAPI. The specificity of anti-V5 antibody was tested with standard western blotting methods.

Immunohistochemical study of brain

To investigate whether DYXC1 is expressed in mature human brain, brain tissue from six deceased individuals was stained with anti-DYXC1-antiserum raised in rabbits against the peptide CATEAKAAAKREDQK (SEQ ID NO:21) (antibody production purchased from Sigma-Genosys). The patients had died of cardiac arrest or ischemic stroke. In the five individuals with stroke, the post-ischemic time before death varied from 15 to 60 h, and brain samples were obtained at rapid autopsies with post-mortem

delays varying from 10 to 40 h. Tissue blocks with cortical and some subcortical tissue were obtained from the core or an area close to the core of the infarction with no specific reference to the topographic location and control samples were dissected from homologous contralateral locations for comparison. Tissues were fixed in formalin and embedded in paraffin, and used for research by permission of the appropriate Ethical Review Board of the Helsinki University Central Hospital. Immunohistochemical methods and use of this post mortem autopsy material for studies on other proteins have been described (25). The dilutions of antiserum used were 1:100-1:200, and all stained sections were compared to adjacent tissue sections incubated with the preimmune serum in identical conditions and dilutions. No antigen retrieval methods were necessary. Light microscopy of tissue sections was performed with Leitz Laborlux D microscope (Leitz, Wetzlar, Germany) equipped with Nikon Coolpix 995 digital camera (Nikon, Japan).

The publications and other materials used herein to illuminate the invention, and in particular, to provide details with respect to its practice, are incorporated herein by reference.

REFERENCES

1. Catts, H.W. (1989) Defining dyslexia as a developmental language disorder. *Ann. Dyslexia,* **39,** 51-64.

2. Grigorenko, E.L. (2001) Developmental dyslexia: An update on genes, brains, and environments. *J. Child. Psychol. Psychiatr.,* **42,** 91-125.

3. Lyon, G.R. (1995) Toward a definition of dyslexia. *Ann. Dyslexia,* **45,** 3-27.

4. Habib, M. (2000) The neurological basis of developmental dyslexia: an overview and working hypothesis. *Brain,* **123 Pt 12,** 2373-99.

5. Paulesu, E., Demonet, J.F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., Cappa, S.F., Cossu, G., Habib, M., Frith, C.D. *et al.* (2001) Dyslexia: cultural diversity and biological unity. *Science,* **291,** 2165-7.

6. Wijsman, E.M., Peterson, D., Leutenegger, A.L., Thomson, J.B., Goddard, K.A., Hsu, L., Berninger, V.W. and Raskind, W.H. (2000) Segregation analysis of phenotypic components of learning disabilities. I. Nonword memory and digit span. *Am. J. Hum. Genet.,* **67,** 631-46.

7. Smith, S.D., Kimberling, W.J., Pennington, B.F. and Lubs, H.A. (1983) Specific reading disability: identification of an inherited form through linkage analysis. *Science,* **219,** 1345-7.

8. Morris, D.W., Robinson, L., Turic, D., Duke, M., Webb, V., Milham, C., Hopkin, E., Pound, K., Fernando, S., Easton, M. *et al.* (2000) Family-based association mapping provides evidence for a gene for reading disability on chromosome 15q. *Hum. Mol. Genet.,* **9,** 843-8.

9. Grigorenko, E.L., Wood, F.B., Meyer, M.S., Hart, L.A., Speed, W.C., Shuster, A. and Pauls, D.L. (1997) Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *Am. J. Hum. Genet.,* **60,** 27-39.

10. Schulte-Körne, G., Grimm, T., Nothen, M.M., Müller-Myhsok, B., Cichon, S., Vogt, I.R., Propping, P. and Remschmidt, H. (1998) Evidence for linkage of spelling disability to chromosome 15. *Am. J. Hum. Genet.,* **63,** 279-82.

11. Cardon, L.R., Smith, S.D., Fulker, D.W., Kimberling, W.J., Pennington, B.F. and DeFries, J.C. (1994) Quantitative trait locus for reading disability on chromosome 6. *Science,* **266,** 276-9.

12. Nopola-Hemmi, J., Taipale, M., Haltia, T., Lehesjoki, A.E., Voutilainen, A. and Kere, J. (2000) Two translocations of chromosome 15q associated with dyslexia. *J. Med. Genet.,* **37,** 771-5.

13. Edelmann, L., Spiteri, E., Koren, K., Pulijaal, V., Bialer, M.G., Shanske, A., Goldberg, R. and Morrow, B.E. (2001) AT-Rich Palindromes Mediate the Constitutional t(11;22) Translocation. *Am. J. Hum. Genet.,* **68,** 1-13.

14. Ramarao, M.K., Bianchetta, M.J., Lanken, J. and Cohen, J.B. (2001) Role of rapsyn tetratricopeptide repeat and coiled-coil domains in self-association and nicotinic acetylcholine receptor clustering. *J. Biol. Chem.,* **276,** 7475-7483.

15. Blatch, G.L. and Lässle, M. (1999) The tetratricopeptide repeat: a structural motif mediating protein- protein interactions. *Bioessays,* **21,** 932-939.

16. Häyrinen, T., Serenius-Sirve, S. and Korkman, M. (1999) *Lukilasse. Lukemisen, kirjoittamisen ja laskemisen seulontatestisto peruskoulun ala-asteen luokille 1-6. Reading and writing test designed for and normated in Finnish elementary school.* Psykologien kustannus Oy, Helsinki.

17. Leinonen, S., Müller, K., Leppänen, P., Aro, M., Ahonen, T. and Lyytinen, H. (2001) Heterogeneity in adult dyslexic readers: Relating processing skills to the speed and accuracy of oral text reading. *Read. Writ. Interdisc. J.,* **14,** 265-296.

18. Wechsler, D. (1992) *Wechsler adult intelligence scale revised (WAIS-R).* Psykologien kustannus Oy and The psychological corporation USA, Helsinki.

19. Wechsler, D. (1984) *Wechsler intelligence scale for children revised (WISC-R).* Psykologien kustannus Oy and The psychological corporation USA, Helsinki.

20. Korkman, M., Kirk, U. and Kemp, S. (1997) *NEPSY. Lasten neuropsykologinen tutkimus. Revised version.* Psykologien kustannus Oy, Helsinki.

21. Denckla, M.B. and Rudel, G.R. (1976) Rapid automatized naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia,* **14,** 471-479.

22. Christensen, A.-L. (1982) *Lurian neuropsykologinen tutkimus. Luria's neuropsychological test.* Psykologien kustannus Oy, Helsinki.

23. Wolf, M. (1986) Rapid alternating stimulus naming in the developmental dyslexias. *Brain Lang.,* **27,** 360-379.

24. Burge, C. and Karlin, S. (1997) Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.,* **268,** 78-94.

25. Lindsberg, P.J., Carpén, O., Paetau, A., Karjalainen-Lindsberg, M.-L. & Kaste, M. (1996) *Circulation* 94, 939-945.

26. Enard, W., Przeworski, M., Fisher, S.E., Lai, C.S.L., Wiebe, V., Kitano, T., Monaco, A.P. & Pääbo, S. (2002) *Nature* 418, 869-872.